



Vers des recommandations plus personnalisées dans les folksonomies

Mohamed Nader Jelassi, Sadok Ben Yahia, Engelbert Mephu Nguifo

► To cite this version:

Mohamed Nader Jelassi, Sadok Ben Yahia, Engelbert Mephu Nguifo. Vers des recommandations plus personnalisées dans les folksonomies. IC - 25èmes Journées francophones d'Ingénierie des Connaissances, May 2014, Clermont-Ferrand, France. pp.187-198. hal-01015545

HAL Id: hal-01015545

<https://inria.hal.science/hal-01015545>

Submitted on 26 Jun 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Vers des recommandations plus personnalisées dans les folksonomies

Mohamed Nader Jelassi^{1,2,3}, Sadok Ben Yahia¹ et Engelbert Mephu Nguifo^{2,3}

¹ Université Tunis El Manar. Faculté des Sciences de Tunis, Tunis, Tunisie.

² Clermont Université, Université Blaise Pascal, LIMOS, BP 10448, F-63000 Clermont-Ferrand, France.

³ CNRS, UMR 6158, LIMOS, F-63171 Aubière, France.

{nader.jelassi@isima.fr, sadok.benyahia@fst.rnu.tn, engelbert.mephu_nguifo@univ-bpclermont.fr}

Résumé :

Plusieurs approches ont été proposées dans la littérature pour personnaliser les recommandations dans les *folksonomies*. Dans ce papier, nous considérons une nouvelle dimension dans les *folksonomies* comme information supplémentaire pour offrir aux utilisateurs une recommandation plus ciblée et mieux conforme à leurs besoins. Cela passe par un regroupement des utilisateurs ayant des intérêts communs sous forme de structures appelées quadri-concepts. Notre approche, dans laquelle nous répondons également au challenge de cold start, est ensuite évaluée sur deux jeux de données du monde réel, MOVIELENS et BOOKCROSSING. Cette évaluation comprend une mesure de la précision et du rappel, une évaluation sociale ainsi que plusieurs métriques d'évaluation comme la diversité, la couverture ou la scalabilité.

Mots-clés : Folksonomie, Personnalisation, Recommandation, Cold Start, Précision

1 Introduction et Motivations

Une *folksonomie* désigne un système de classification collaborative par les internautes (Mika (2007)). L'idée est de permettre à des utilisateurs de partager et de décrire des objets via des mots-clés (tags) librement choisis. Formellement, une *folksonomie* est composée de trois ensembles : un ensemble \mathcal{U} d'utilisateurs, un ensemble \mathcal{T} de tags (ou étiquettes) et un ensemble \mathcal{R} de ressources (films, livres, sites web, photos, etc.). Les utilisateurs sont les acteurs principaux du système et contribuent au contenu par l'ajout de ressources et l'affectation de tags. Cependant, il s'avère que le choix de tags et de ressources partagées par un utilisateur d'une *folksonomie* varie selon plusieurs critères : le genre, l'âge ou encore la profession de celui qui partage l'information. Ainsi, les *folksonomies* doivent tenir compte de telles informations lors de la recommandation de tags ou de ressources. Cela a incité les chercheurs à proposer des systèmes de recommandation personnalisés afin de suggérer les tags et ressources les plus appropriés aux utilisateurs et de répondre aux besoins de chaque utilisateur. En effet, le domaine de personnalisation tente de fournir des solutions afin d'aider les utilisateurs à partager les bons tags et les bonnes ressources parmi le très grand nombre de données dans les *folksonomies*. Ainsi, un système de recommandation offre à l'utilisateur une liste de tags ou de ressources recommandés qu'il est susceptible d'aimer et lui permet de trouver plus facilement ses tags et ressources préférés dans la *folksonomie* (Ricci *et al.* (2011)). De plus, la personnalisation tente d'aider les utilisateurs à aborder le problème de surcharge d'information (Das *et al.* (2012)). Et pour réussir ou tenter de répondre au mieux aux attentes de chaque utilisateur de la *folksonomie*, il est utile d'avoir plus d'informations sur lui. En effet, son âge, sa profession ou sa localisation sont des informations qui sont susceptibles de nous aider dans le processus de personnalisation

de recommandation. Pour atteindre cet objectif, nous considérons une nouvelle dimension dans une *folksonomie*, classiquement composée de trois dimensions (utilisateurs, tags et ressources), et nous proposons une approche de regroupement des utilisateurs aux intérêts équivalents sous forme de structure appelées concepts quadratiques (Jelassi *et al.* (2013)). Cette quatrième dimension peut recouvrir différents aspects : par exemple le profil (genre, âge, profession, *ldots*) comme mentionné ci-dessus, ou le temps si on veut étudier la dynamique temporelle des *folksonomies*. Dans ce papier, nous traitons la quatrième dimension de manière indifférente pour l'aspect méthodologique, mais afin de comparer notre méthode avec les travaux de la littérature, nous focaliserons sur l'aspect profil.

Une question se pose alors : pourquoi les concepts quadratiques ? D'un côté, si on peut facilement étudier les tags utilisés par un seul utilisateur sur une ressource, il est évident de constater que la tâche devient rapidement intraitable lorsque cela implique plusieurs utilisateurs et plusieurs ressources. D'un autre côté, les tags (ou ressources) recommandés s'avèrent ne pas être très spécifiques (Jäschke *et al.* (2007)), *i.e.*, des tags qui sont des mots "bateau" ou bien des ressources vagues ne correspondant pas aux besoins spécifiques de l'utilisateur. Grâce aux concepts quadratiques, nous pouvons résoudre ces problèmes. En effet, d'un côté, les concepts quadratiques sont des structures regroupant les tags et ressources en commun à un ensemble maximal d'utilisateurs. D'un autre côté, dans un concept quadratique, les tags et ressources qui ont été utilisés en combinaison seront regroupés d'où un résultat plus spécifique et répondant au besoin de notre système de recommandation. Ces concepts sont une représentation réduite de la *folksonomie* qui peut contenir des milliers de quadruplets dans la vraie vie. Une fois extraits, ces quadri-concepts sont utilisés pour notre algorithme de recommandation personnalisée multi-mode (utilisateurs, tags et ressources). Nous menons ensuite une évaluation étendue de notre système de recommandation sur deux jeux de données : tout d'abord, nous calculons la précision et le rappel de notre système, ensuite nous analysons ses différentes propriétés comme la diversité ou la scalabilité et nous menons une évaluation sociale sur différents utilisateurs. De plus, nous étudions la capacité de notre système à proposer des recommandations aux nouveaux utilisateurs, *i.e.*, le problème de cold start (Lam *et al.* (2008)).

Le reste du papier est organisé comme suit : nous étudions les principales approches de la littérature dans la Section 2. Dans la Section 3, nous proposons un système personnalisé de recommandation. Nous menons une étude expérimentale dans la Section 4. Enfin, nous concluons notre papier avec des perspectives pour nos travaux futurs dans la Section 5.

2 Travaux connexes

Dans un souci d'améliorer les recommandations dans les *folksonomies*, plusieurs travaux ont été proposés dans la littérature. Dans (Diederich & Iofciu (2006)), les auteurs utilisent la "*personomie*" d'un utilisateur, *i.e.*, les tags qui lui sont relatifs, afin de lui recommander des utilisateurs ayant partagé des tags et ressources similaires. Tout d'abord, ils construisent un profil pour chaque utilisateur. Ensuite, à partir de ce profil, les auteurs sont capables de recommander des utilisateurs (dits *collaborateurs*) en utilisant une mesure de similarité entre utilisateurs. Cette mesure, qui s'appuie uniquement sur les tags utilisés par les utilisateurs, n'offre pas une information complète sur les utilisateurs. Plus récemment, dans (Hu *et al.* (2011)), les auteurs se basent à la fois sur l'historique de tagging (tags et ressources) des utilisateurs et sur leurs contacts sociaux. La limite de cette approche est qu'elle requiert qu'un utilisateur doit posséder

des contacts sociaux afin d'avoir des recommandations de tags. Dans (Jäschke *et al.* (2007)), Hotho *et al.* ont proposé des recommandations de tags dans les *folksonomies* basées sur les tags les plus utilisés. Cependant, ces recommandations ne sont absolument pas personnalisées étant donné que les mêmes tags sont proposés à chaque utilisateur. Lipczak a proposé dans (Lipczak (2008)) un système de recommandation de tags en trois étapes. À partir des tags annotés aux ressources, l'auteur ajoute des tags proposés par un lexique basé sur les co-occurrences de tags sur les mêmes ressources. Ensuite, le système filtre les tags déjà utilisés par l'utilisateur. Toutefois, malgré cette étape de filtrage, la recommandation ne paraît pas être personnalisée étant donné qu'elle cherche des tags co-occurrençant sur d'autres annotations. L'approche revient ensuite à enlever les tags précédemment annotés par l'utilisateur de ceux qui sont suggérés. Dans (Landia & Anand (2009)), les auteurs ont proposé une nouvelle approche combinant la similarité à la fois entre ressources et entre utilisateurs afin de recommander des tags personnalisés. En effet, deux utilisateurs sont considérés comme similaires s'ils ont assigné les mêmes tags aux mêmes ressources. Toutefois, il est rare de trouver pareille situation dans des *folksonomies* où les tags utilisés par deux utilisateurs sur les mêmes ressources sont identiques.

Dans notre approche, nous insistons sur le nécessaire recours à des informations supplémentaires et à les combiner à l'historique de tagging afin d'améliorer les recommandations. Toutes ces informations seront représentées par des quadri-concepts. Ainsi, dans ces structures, nous nous focalisons non seulement sur les tags/ressources les plus utilisés, mais également sur ceux qui ont été utilisés en combinaison, obtenant ainsi un résultat plus spécifique. Contrairement aux approches de la littérature qui se limitent à l'information $\langle \text{utilisateur}, \text{tag}, \text{ressource} \rangle$, nous étendons ce triplet par l'information contenue dans la quatrième dimension.

3 Un nouveau système de recommandation personnalisée

Nous commençons par présenter une extension de la notion de *folksonomie* (Jäschke *et al.* (2008)) par l'ajout d'une quatrième dimension.

Une **v-folksonomie** est un ensemble de tuples $\mathcal{F}_v = (\mathcal{U}, \mathcal{T}, \mathcal{R}, \mathcal{V}, Y)$ où $\mathcal{U}, \mathcal{T}, \mathcal{R}$ et \mathcal{V} sont des ensembles finis dont les éléments sont appelés **utilisateurs**, **tags**, **ressources** et **variables**. $Y \subseteq \mathcal{U} \times \mathcal{T} \times \mathcal{R} \times \mathcal{V}$ représente une relation quadratique où chaque élément $y \subseteq Y$ peut être représenté par un quadruplet : $y = \{(u, t, r, v) \mid u \in \mathcal{U}, t \in \mathcal{T}, r \in \mathcal{R}, v \in \mathcal{V}\}$ ce qui veut dire que l'utilisateur u a annoté la ressource r via le tag t à travers la variable v . Nous considérons que deux utilisateurs sont *proches* s'ils partagent au moins une même variable en commun.

Nous définissons maintenant un concept quadratique (Jelassi *et al.* (2013)).

Un **concept quadratique** (ou quadri-concept) d'une *v-folksonomie* $\mathcal{F}_v = (\mathcal{U}, \mathcal{T}, \mathcal{R}, \mathcal{V}, Y)$ est un quadruplet (U, T, R, V) avec $U \subseteq \mathcal{U}, T \subseteq \mathcal{T}, R \subseteq \mathcal{R}$ et $V \subseteq \mathcal{V}$ avec $U \times T \times R \times V \subseteq Y$ tel que le quadruplet (U, T, R, V) est maximal. Un quadri-concept est donc la version quadri-dimensionnelle d'un ensemble fermé. Par ailleurs, la quatrième dimension \mathcal{V} peut recouvrir différents aspects (*e.g.*, le profil, le temps). Dans ce papier, nous traitons la quatrième dimension de manière indifférente pour l'aspect méthodologique, mais afin de comparer notre méthode avec les travaux de la littérature, nous focaliserons sur l'aspect profil.

Afin de permettre l'extraction de l'ensemble de quadri-concepts fréquents à partir d'une *v-folksonomie* donnée, nous pouvons utiliser l'un des deux algorithmes de la littérature dédiés à cette tâche : QUADRICONS (Jelassi *et al.* (2013)) ou DATAPEELER (Cerf *et al.* (2009)). Les deux algorithmes prennent en entrée une *v-folksonomie* ainsi que quatre seuils minimaux de

support (un pour chaque dimension) et donnent en sortie l'ensemble de **quadri-concepts** vérifiant ces seuils. Un quadri-concept **fréquent** est un quadri-concept dont chaque ensemble (utilisateur, tag, ressource et variable) a une cardinalité supérieure ou égale à son seuil de support correspondant. Par ailleurs, il est important de noter que, même si l'étape d'extraction des quadri-concepts est une phase qui peut avoir une complexité exponentielle, elle se passe hors-ligne et n'est exécutée qu'une seule fois. En effet, notre algorithme PERSOREC parcourt des quadri-concepts déjà extraits. Ainsi, notre système de recommandation ne souffre pas du coût d'extraction des quadri-concepts à chaque recommandation. Cette phase peut donc être vue comme un pré-traitement à la phase de recommandation. Dans ce qui suit, nous introduisons notre algorithme de recommandation personnalisée pour les *folksonomies* (Algorithme 1) qui donne en sortie trois différents ensembles : un ensemble d'utilisateurs proposés, un ensemble de tags suggérés et un ensemble de ressources recommandées.

Algorithme 1 : PERSOREC

Données : l'ensemble des quadri-concepts fréquents QC , un utilisateur u avec sa variable v et une ressource r .

Résultats : l'ensemble d'utilisateurs proposés \mathcal{PU} , l'ensemble des tags suggérés \mathcal{ST} et l'ensemble des ressources recommandées \mathcal{RR} .

```

1  début
2  |   pour chaque quadri-concept  $qc \in QC$  faire
3  |   |   si  $v \in qc.VARIABLES$  alors
4  |   |   |   si  $u \notin qc.UTILISATEURS$  alors
5  |   |   |   |   /*Proposition d'utilisateurs */
6  |   |   |   |    $\mathcal{PU} = \mathcal{PU} \cup qc.extent$  ;
7  |   |   |   |   /*Suggestion de Tags*/
8  |   |   |   |   si  $r \in qc.RESSOURCES$  alors
9  |   |   |   |   |    $\mathcal{ST} = \mathcal{ST} \cup qc.Tags$  ;
10 |   |   |   |   /*Recommandation de Ressources */
11 |   |   |   |    $\mathcal{RR} = \mathcal{RR} \cup qc.RESSOURCES$  ;
12 |   retourner  $(\mathcal{PU}, \mathcal{ST}, \mathcal{RR})$  ;
13 fin
```

PERSOREC opère comme suit : dans la Ligne 3, il parcourt l'ensemble des quadri-concepts fréquents en cherchant ceux dont les utilisateurs sont proches de u selon la variable v . Le test de la Ligne 4 permet de filtrer les tags et ressources déjà partagés par l'utilisateur u ; cette stratégie est inspirée par celle de (Lipczak (2008)). Ensuite, pour chaque tâche, PERSOREC fonctionne comme suit : pour la tâche de *Proposition d'utilisateurs* (Ligne 6), c'est la partie *utilisateurs* du quadri-concept qc qui est ajoutée à l'ensemble \mathcal{PU} des utilisateurs proposés. Cette tâche aide à connecter les utilisateurs qui ont des intérêts communs et aide également à promouvoir le partage de ressources. Pour la tâche de *Suggestion de tags* (Lignes 8 et 9), le but est de suggérer des tags personnalisés à un utilisateur qui souhaite ajouter une ressource à la *folksonomie*. Cette tâche a plusieurs avantages : elle rappelle à l'utilisateur ce dont une ressource s'agit, accroît l'annotation des ressources et permet de consolider le vocabulaire des utilisateurs

(Ricci *et al.* (2011)). Pour cette tâche, nous ajoutons donc les tags affectés à la ressource r par les utilisateurs proches de u à l'ensemble ST . Quant à la tâche de *Recommandation de ressources* (Ligne 12), le but est de proposer une liste personnalisée de ressources conforme aux intérêts de l'utilisateur u ; ces ressources sont ajoutées à l'ensemble \mathcal{RR} . Dans ce qui suit, nous évaluons notre approche sur deux jeux de données standard pour la recommandation..

4 Résultats et Discussion

Les deux jeux de données du monde réel utilisés pour notre évaluations sont : tout d'abord, le jeu de données filmographique MOVIELENS (<http://movielens.umn.edu/>) qui est un système de recommandation et un site web communautaire qui permet aux utilisateurs de partager des films en les annotant par des tags. Le jeu de données, utilisé pour nos expérimentations, est téléchargeable gratuitement (<http://www.grouplens.org/node/73>) et contient 95580 tags appliqués à 10681 films par 71567 utilisateurs (*e.g.*, $\langle \text{Alex}, X\text{-Files}, \text{sciencefiction} \rangle$). Le second jeu de données utilisé est BOOKCROSSING (<http://www.bookcrossing.com/>) qui est un "club de lecture" en ligne dont le but est de "faire du monde entier une bibliothèque". Contrairement à MOVIELENS, les utilisateurs de BOOKCROSSING n'utilisent pas les tags pour annoter les ressources, *i.e.*, les livres, mais plutôt sur les notes. Ainsi, les utilisateurs choisissent une note entre 1 et 10, *i.e.*, plus la note est élevée, plus l'appréciation est meilleure. Le jeu de données utilisé est téléchargeable gratuitement (<http://www.grouplens.org/node/74>) et contient 278858 utilisateurs, 1149780 notes et 271379 livres (*e.g.*, $\langle \text{Regina}, \text{DaVinciCode}, 9 \rangle$). Pour les besoins de comparaison avec les approches de la littérature, nous avons choisi, dans ce qui suit, le profil des utilisateurs pour modéliser la variable v . Ainsi, nous considérons à présent que deux utilisateurs sont *proches* s'ils partagent au moins une même information de profil en commun (*e.g.*, un même âge, une même profession, etc.). À cet effet, des informations supplémentaires sur les utilisateurs sont disponibles dans les deux jeux de données et forment le profil des utilisateurs (la quatrième dimension d'une p -folksonomie) et qui renseigne, pour MOVIELENS, sur le **genre** de l'utilisateur (masculin ou féminin) et sa **profession** (au nombre de 21, qui peut être éducateur, écrivain, étudiant, scientifique, etc.). Pour BOOKCROSSING, nous avons des informations sur la localisation des utilisateurs. Par ailleurs, les deux jeux de données fournissent des informations sur l'**âge** des utilisateurs qui est divisé en cinq tranches : (i) 7 – 18 ans ; (ii) 19 – 24 ans ; (iii) 25 – 35 ans ; (iv) 36 – 45 ans et (v) 46 – 73 ans.

4.1 Evaluation des recommandations : Precision et Rappel de PERSOREC

Chacun des jeux de données MOVIELENS et BOOKCROSSING a été partitionné en deux échantillons : un échantillon contenant 80% des utilisateurs a été utilisé comme **base d'apprentissage** et un échantillon contenant les 20% d'utilisateurs restants, a été utilisé pour la validation de nos tests (*i.e.*, **base de test**). Pour chaque utilisateur du deuxième échantillon (*i.e.*, utilisateur test), 20% aléatoires de ses tags et ressources sont considérées comme ensemble de test/réponse et 80% comme son ensemble d'apprentissage. Pour chaque utilisateur test, notre algorithme de recommandation génère une liste d'éléments (utilisateurs, tags ou ressources) en se basant sur son ensemble d'apprentissage. Si un élément de la liste de recommandation se trouve également dans l'ensemble de test de cet utilisateur, alors l'élément est considéré comme **pertinent**.

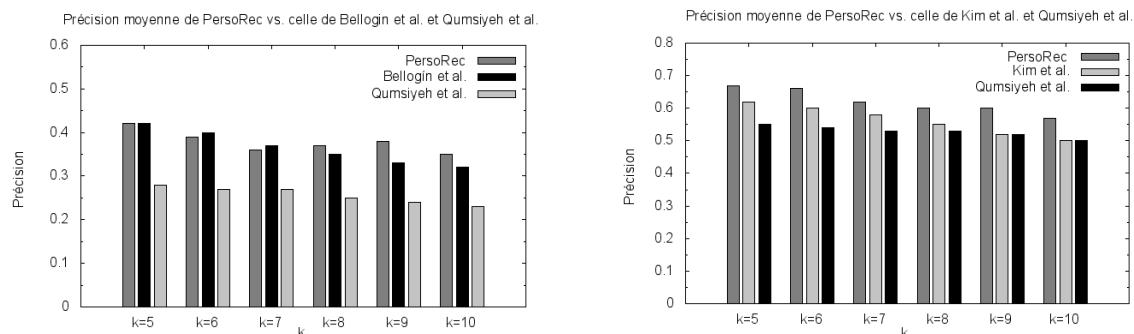


FIGURE 1 – Précision de PERSOREC vs. les autres approches pour (à gauche) la recommandation de films et (à droite) la recommandation de livres.

Évaluer l'efficacité d'un algorithme de recommandation est loin d'être trivial. En premier lieu, parce que différents algorithmes peuvent être meilleurs ou moins bons en fonction du jeu de données sur lequel ils sont appliqués (Herlocker *et al.* (2004)). Néanmoins, pour déterminer l'efficacité d'un système, nous pourrions appliquer les métriques classiques de recherche d'informations : la précision et le rappel (Baeza-Yates & Ribeiro-Neto (1999)). Ces mesures représentent la qualité de la recommandation, c'est-à-dire à quel point les suggestions proposées sont conformes aux intérêts de l'utilisateur. Tout d'abord, la précision détermine la probabilité qu'un élément recommandé soit pertinent. Dans nos expérimentations, nous avons également fait varier le nombre de recommandations proposées dont l'utilisateur peut spécifier le nombre k de réponses les plus pertinentes que le système doit lui retourner. Cela permet surtout d'éviter de submerger l'utilisateur par un grand nombre de réponses en lui retournant que le nombre de réponses les plus pertinentes qu'il souhaite. Dans ce qui suit, nous nous intéressons à la tâche de recommandation de ressources et nous évaluons la précision de notre approche vs. les travaux pionniers qui ont un objectif commun avec la nôtre, *i.e.*, ceux de Bellogin *et al.* (Bellogín *et al.* (2013)), Qumsiyeh *et al.* (Qumsiyeh & Ng (2012)) et Kim *et al.* (Kim *et al.* (2011)). Ces approches n'utilisent pas de quatrième dimension mais font appel au profil des utilisateurs comme information complémentaire pour la tâche de recommandation.

Ainsi, la Figure 1 montre les différentes valeurs de précision obtenues par notre algorithme de recommandation vs. ses concurrents pour différentes valeurs de k variant entre 5 et 10 sur les deux jeux de données. En général, nos recommandations pour les utilisateurs de MOVIELENS et BOOKCROSSING répondent à leurs attentes. En effet, les recommandations sont pertinentes à 38% et 62%, respectivement, pour MOVIELENS et BOOKCROSSING, ce qui surpasse, pour la plupart des cas, la précision de ses concurrents. Ainsi, pour le jeu de données MOVIELENS, si notre approche fait jeu égal avec celle de Bellogin *et al.* (une précision tout juste meilleure de 2%), l'écart avec l'approche de Qumsiyeh *et al.* est bien plus considérable (une précision plus élevée de 38%). Quant au jeu de données BOOKCROSSING, notre précision est plus grande de, respectivement, 27% et 29% que celles de Kim *et al.* et Qumsiyeh *et al.*. Par ailleurs, les résultats ont montré que PERSOREC atteint ses meilleures performances lorsque la valeur de k est égale à 5. Cela est dû au fait que les cinq premières recommandations correspondent aux besoins des utilisateurs et que lorsque le nombre de recommandations augmente, cela entraîne inévitablement une diminution de la précision étant donné que l'utilisateur choisit moins de

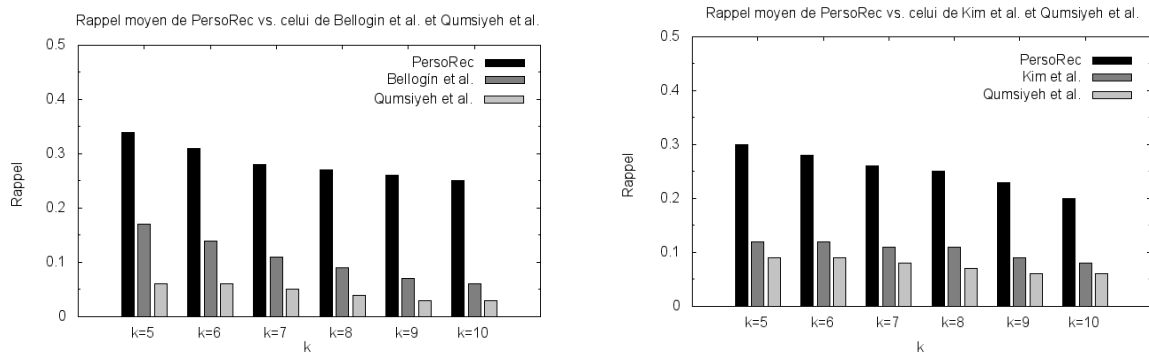


FIGURE 2 – Rappel de PERSOREC vs. les autres approches pour (à gauche) la recommandation de films et (à droite) la recommandation de livres.

ressources que celles qui lui sont recommandées. Quant à la différence entre notre précision et celle des autres approches, nous l'expliquons par le fait que l'utilisation des quadri-concepts améliore les recommandations en suggérant les tags et ressources les plus proches des besoins des utilisateurs. En effet, alors que les travaux connexes se concentrent sur les éléments les plus populaires (livres, films, tags), les quadri-concepts offrent à nos utilisateurs, les tags et les ressources qui ont été partagés en commun par des utilisateurs aux profils proches.

Quant au rappel, il mesure le pourcentage de recommandations pertinentes retournées à l'utilisateur parmi l'ensemble total de recommandations pertinentes. La Figure 2 démontre les différentes valeurs de rappel obtenues par PERSOREC vs. les autres approches de la littérature pour différentes valeurs de k allant de 5 à 10 sur les jeux de données MOVIELENS et BOOKCROSSING. Les résultats montrent que notre algorithme surpasse nettement les approches de la littérature. En effet, PERSOREC a une valeur moyenne de rappel égale à 30% sur MOVIELENS vs. 17% pour Bellogin *et al.* et 6% pour Qumsiyeh *et al.*. Quant au jeu de données BOOKCROSSING, notre rappel est, respectivement, 2,50 et 2,77 fois meilleur que celui de Kim *et al.* et Qumsiyeh *et al.*. Cette différence de performances démontre que sur l'ensemble total des éléments pertinents, PERSOREC est capable d'en recommander, à ses utilisateurs, une portion plus grande que ses concurrents. Grâce aux quadri-concepts, qui représentent des structures représentatives de la *folksonomie*, PERSOREC recommande donc des éléments partagés par des utilisateurs qui sont susceptibles d'être ensuite partagés par des utilisateurs avec un profil proche.

Le *F1-Score* (ou *mesure F1*) combine à la fois la précision et le rappel qui peut être interprétée comme une moyenne pondérée de ces deux mesures Herlocker *et al.* (2004). La meilleure valeur du *F1-Score* est égale à 1 tandis que la pire valeur correspond à 0. Dans le domaine de la recommandation, cette mesure est un bon indicateur de l'utilité des recommandations. La Figure 3 démontre les différentes valeurs de *F1-Score* obtenues par PERSOREC vs. les autres approches de la littérature pour différentes valeurs de k allant de 5 à 10 sur les datasets MOVIELENS et BOOKCROSSING. Les résultats montrent très logiquement que PERSOREC surpasse ses concurrents sur les deux datasets puisque les valeurs de rappel et de précision de PERSOREC sont supérieures à celles des autres approches.

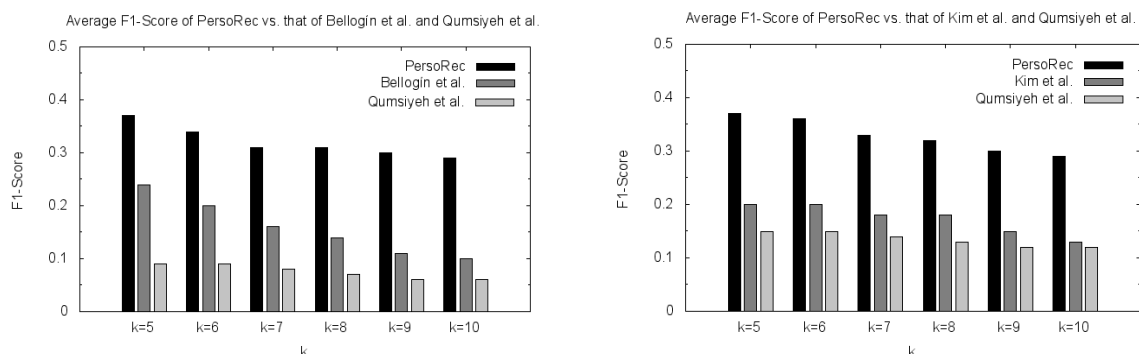


FIGURE 3 – F1-Score de PERSOREC vs. les autres approches pour (à gauche) la recommandation de films et (à droite) la recommandation de livres.

4.2 Évaluation Sociale

Dans ce qui suit, nous étudions l'évaluation sociale de notre système de recommandation. Nous analysons ce qui se passe **après** l'étape de recommandation, *i.e.*, si l'utilisateur cible a vraiment aimé les recommandations et si les utilisateurs (amis) qu'on lui a proposés adoptent le même comportement social. Pour ce faire, nous étudions trois différents cas réels de recommandation avec BOOKCROSSING et un cas réel pour MOVIELENS.

Ainsi, pour le premier jeu de données, nous avons choisi trois utilisateurs avec différents âges et pays : *skinner* (38 ans, New York, USA), *herge* (26 ans, Seixal, Portugal) et *benjamin* (15 ans, Texas, USA). En premier lieu, notre algorithme recommande au premier utilisateur trois différents livres de la franchise *Harry Potter* ainsi que quatre nouveaux amis : *ross* (43 ans, Illinois, USA), *fran* (54 ans, California, USA), *emma* (40 ans, Oregon, USA) et *anna_lucia* (36 ans, Teheran, Iran). Il s'est avéré plus tard que ces nouveaux amis ont également partagé tous les livres de la franchise *Harry Potter*. De plus, *skinner*, ainsi que ses amis recommandés, ont attribué aux livres recommandés la note de 9 ce qui démontre une réelle appréciation des recommandations. En second lieu, nous avons recommandé à *herge* trois différents livres (*Da Vinci Code*, *Wild Animus* et *The Joy Luck Club*) ainsi que quatre amis âgés entre 25 et 35 ans : trois d'entre eux venant des USA (*Kansas*, *Wisconsin* et *Virginia*) et le quatrième du Canada (*Ottawa*). Cependant, bien qu'il ait partagé les trois livres, *herge* ne les a pas notés, et parmi ses "nouveaux amis", l'un d'eux fut vraiment intéressé par les mêmes livres. En dernier lieu, notre algorithme a généré pour *benjamin* deux livres (*Harry Potter and the Prisonnier of Azkaban* et *Harry Potter and Cup of Fire*) ainsi qu'un nouvel ami, *i.e.*, *baefire* (12 ans, Illinois, USA). Il se trouve, qu'après la recommandation, les deux utilisateurs ont partagé les livres en leur attribuant la note maximale (10) ce qui indique qu'ils ont vraiment bien apprécié les livres recommandés.

Quant au jeu de données MOVIELENS, notre utilisateur cible est *Bruce* (47 ans, Homme, Educateur). PERSOREC lui a recommandé quatre films : *Star Wars*, *The Return of the Jedi*, *God Father 1 and 2* ainsi que deux nouveaux utilisateurs : *Slioua* (49 ans, Femme, Educatrice) et *Nina* (49 ans, Homme, Educateur). Tout d'abord, nous pouvons remarquer que *Bruce* a apprécié les films recommandés en leur attribuant la note maximale de 5. Ensuite, nous avons noté que ses amis recommandés ont aussi partagé les mêmes films avec une note moyenne de 4 ce qui démontre que Bruce et ses nouveaux amis ont des intérêts communs pour les mêmes films.

Dans ce qui suit, nous proposons plusieurs métriques afin d'évaluer les propriétés de notre système de recommandation. Ces métriques sont définies comme la capacité d'un système de recommandation à suggérer à l'utilisateur des éléments pertinents mais non populaires.

4.3 Propriétés de PERSOREC

Bien qu'elle soit une tâche cruciale, la recommandation d'utilisateurs, de tags et de ressources s'avère parfois insuffisante pour déployer un bon système de recommandation. Souvent, les utilisateurs peuvent être intéressés par plus qu'une bonne recommandation : découvrir de nouveaux éléments, la diversité des éléments, etc. Ainsi, nous devons identifier l'ensemble de propriétés qui influent sur la réussite d'un système de recommandation (Ricci *et al.* (2011)) :

Couverture de l'Espace Utilisateur. La couverture de l'espace utilisateur est définie comme étant la portion d'utilisateurs pour laquelle le système peut recommander des éléments. Les algorithmes capables de fournir des recommandations à la majorité des utilisateurs sont donc particulièrement appréciés. Ceci dit, PERSOREC est capable de donner des recommandations à tous les utilisateurs de la *folksonomie* indépendamment du fait qu'un utilisateur ait tagué moins d'éléments qu'un seuil défini ou le fait qu'il doit avoir un certain nombre d'amis. Dès qu'un utilisateur est ajouté à la *folksonomie*, ses informations personnelles sont suffisantes pour obtenir des recommandations de tags, de ressources et d'utilisateurs (comme démontré dans le critère suivant). Nous avons, par ailleurs, calculé le pourcentage de profils couverts par PERSOREC, ce qui a donné les résultats suivants : 100% de genres (homme et femme), 100% des catégories d'âge, 100% des métiers et 88% des pays¹.

Problème du cold start. Un problème récurrent dans le processus de recommandation est celui du "démarrage à froid" ou cold start (Ricci *et al.* (2011)), *i.e.*, la performance du système vis-à-vis des nouveaux utilisateurs. Dans ce qui suit, nous tentons de répondre à ce problème. Un utilisateur est considéré comme nouveau s'il n'a encore tagué aucune ressource. Contrairement à la majorité des approches de la littérature, PERSOREC ne requiert pas qu'un utilisateur ait partagé un nombre minimum de ressources avant d'être considéré par le système de recommandation. Cela revient au fait que PERSOREC regarde d'abord le profil d'un utilisateur, *i.e.*, ses informations personnelles (âge, profession, etc.) pour lui fournir des recommandations. Ensuite, lorsque ce même utilisateur commence à annoter des ressources avec des tags, le système pourra lui fournir de nouvelles recommandations selon ce qu'il a partagé. Avec cette extension, PERSOREC est maintenant capable de recommander à ses utilisateurs les nouvelles ressources ajoutées et de prendre en compte les nouveaux tags et utilisateurs sans avoir à redémarrer le processus d'extraction des quadri-concepts. Ainsi, notre méthode peut devenir incrémentale.

Diversité. Recommander un ensemble d'éléments qui sont similaires n'est pas aussi utile pour les utilisateurs, qui préfèrent la diversité, *i.e.*, des recommandations qui sont différentes et *distantes*. Par exemple, un utilisateur préférera une recommandation de cinq livres écrits par cinq auteurs différents à une recommandation de cinq livres écrits par un même auteur. Afin de mesurer la diversité de nos recommandations sur le jeu de données BOOKCROSSING, nous utilisons la métrique de distance d (*cf.*, équation 1) qui mesure la distance entre l'élément recommandé et un ensemble d'éléments déjà tagués par l'utilisateur. La métrique de distance d est définie comme suit (Ricci *et al.* (2011)) :

1. À partir de l'ensemble des 13625 pays représentés dans BOOKCROSSING, nous avons évalué la couverture de PERSOREC sur les pays les plus représentés, *i.e.*, les pays présents dans, au moins, 500 quadruplets.

$$d(b, B) = \frac{1 + C_B - C_{B.w(b)}}{1 + C_B} \quad (1)$$

où b est le livre recommandé, B l'ensemble de livres déjà lus par l'utilisateur ciblé, C_B le nombre maximal de livres écrits par un même auteur dans l'ensemble B et $C_{B.w(b)}$ le nombre de livres écrits par l'auteur de b dans l'ensemble B . À noter que la valeur de d se tient dans l'intervalle unitaire. Nous utilisons cette métrique de la manière suivante : nous calculons d'abord la distance entre chaque livre recommandé et le reste de la liste des livres recommandés et ensuite, nous calculons la moyenne de ces résultats afin d'obtenir le score de diversité. Nous avons donc obtenu un score de diversité égal à 0,56 pour l'ensemble des trois utilisateurs (de l'évaluation sociale) avec un score maximal égal à 1 atteint pour le second utilisateur. Ce dernier score s'explique par le fait que nous avons recommandé au deuxième utilisateur trois livres d'auteurs différents, ce qui, en plus de le surprendre, plaît à un utilisateur intéressé par des recommandations diverses.

Scalabilité. L'approche standard pour évaluer la scalabilité d'un système est d'évaluer la complexité de l'algorithme dédié en termes de temps d'exécution ou/et de mémoire requise. Ainsi, nous calculons le temps moyen d'exécution (en millisecondes) de nos recommandations sur les deux jeux de données pour les tâches de recommandation de ressources (dénnotée Tâche 1) et de proposition d'utilisateurs (dénnotée Tâche 2). Le Tableau 1 affiche tout d'abord le temps d'exécution de PERSOREC sur le jeu de données MOVIELENS qui contient 100000 quadruplets $\langle \text{utilisateur}, \text{tag}, \text{ressource}, \text{profil} \rangle$. Chaque quadri-concept extrait contient au moins un tag, une ressource et une information de profil, alors que nous faisons varier le support minimum d'utilisateurs (minsupp_u), *i.e.*, le nombre minimum d'utilisateurs par quadri-concept. Par exemple, lorsque minsupp_u est égal à 6, nous avons 13461 quadri-concepts où chaque concept contient, au moins, 6 utilisateurs. Le Tableau 1 démontre les bonnes performances de PERSOREC pour toutes les valeurs de minsupp_u . Tandis que le nombre de quadri-concepts augmente rapidement (de 221 à 13461), le temps d'exécution des recommandations générées par PERSOREC est en moyenne de 2 ms et de 8 secondes pour, respectivement, la première et deuxième tâche. Le nombre total des recommandations (*i.e.*, le nombre d'utilisateurs uniques) est égal à 865 pour la plus petite valeur de minsupp_u . Le Tableau 1 affiche également les performances de PERSOREC sur le jeu de données BOOKCROSSING qui contient 762000 quadruplets. Nous avons fait varier le nombre minimum d'utilisateurs par quadri-concept de 30 à 10. Tout d'abord, nous pouvons voir que le nombre maximal de quadri-concepts extraits est égal à 13461 ce qui représente seulement 1,76% de la *folksonomie* ; ce qui démontre, une fois de plus, l'utilité des quadri-concepts, qui sont une représentation réduite de la *folksonomie*. Contrairement à MOVIELENS, le nombre d'utilisateurs uniques, *i.e.*, le nombre total de recommandations, augmente considérablement tandis que le nombre de quadri-concepts fréquents croît légèrement (en raison des valeurs élevées de minsupp_u). Cependant, si PERSOREC affiche toujours des bonnes performances pour la tâche de recommandation de ressources où le temps moyen d'exécution est de 28 ms, la tâche de proposition d'utilisateurs devient légèrement plus lente étant donné que chaque quadri-concept contient au moins 10 utilisateurs. Cependant, le temps d'exécution demeure raisonnable à hauteur de 384 secondes en moyenne.

Comparaisons avec les travaux de la littérature. Afin de mettre en lumière les plus-values de notre approche par rapport à ses prédécesseurs, nous surlignons dans le Tableau 2 les différents critères (discutés plus en détails dans la Section 4.3) qu'un système de recommandation

	<i>minsupp_u</i>	$ QC $	# Utilisateurs Uniques	Tâche 1 (ms)	Tâche 2 (ms)
(MOVIELENS)	20	221	526	0, 1	2, 6
	16	500	605	0, 2	3, 9
	12	1295	668	0, 7	6, 1
	8	5123	805	4, 0	13, 5
	6	13461	865	12, 7	23, 3
(BOOKCROSSING)	30	553	6789	0, 9	149, 8
	20	1486	9092	4, 9	296, 9
	16	2638	10397	13, 0	415, 5
	12	5698	12239	45, 0	542, 3
	10	10100	13457	114, 7	586, 8

TABLE 1 – Temps moyen d'exécution des recommandations de PERSOREC.

doit vérifier Ricci *et al.* (2011). Le point d'interrogation (" ? ") dénote qu'une information est manquante dans l'approche et est difficile à vérifier. Nous pouvons remarquer, par exemple, qu'aucune des approches n'offre une recommandations multi-mode (d'utilisateurs, tags et de ressources en même temps). Comme il sera démontré dans la Section 4.3, notre approche diffère de ses prédécesseurs en tenant en compte les nouveaux utilisateurs (les critères *couverture* et *cold start*) en leur fournissant des recommandations sans qu'ils n'aient déjà tagué par le passé. Par ailleurs, si la plupart des approches satisfont le critère de *diversité*, d'autres critères comme la *scalabilité* (le passage à l'échelle) sont difficiles à vérifier étant donné que leurs auteurs n'ont pas donné suffisamment d'informations sur leurs approches.

	Multi-Mode	Couverture	Cold Start	Diversité	Scalabilité
Diederich & Iofciu (2006)	Non	Non	Non	Oui	Non
Basile <i>et al.</i> (2007)	Non	Non	Non	?	?
Jäschke <i>et al.</i> (2007)	Non	Non	Non	Oui	?
Lipczak (2008)	Non	Non	Non	Oui	?
Landia & Anand (2009)	Non	Non	Non	Non	?
De Meo <i>et al.</i> (2010)	Non	Non	Non	?	?
Hu <i>et al.</i> (2011)	Non	Non	Non	Oui	?
Notre approche	Oui	Oui	Oui	Oui	Oui

TABLE 2 – Les différentes approches de la littérature en bref.

5 Conclusion et Perspectives

Dans ce papier, nous avons introduit une nouvelle approche de recommandation basée sur les concepts quadratiques, et permettant d'offrir un choix personnalisé de tags et de ressources aux utilisateurs. L'évaluation de notre algorithme de recommandation a donné de bons résultats selon les métriques introduites. Nous avons également mené une étude de cas sur six sujets différents afin d'avoir leur feedback et dont les résultats sont disponibles dans la version étendue de notre papier (Jelassi *et al.* (2014)). Parmi les perspectives de nos travaux, nous pouvons considérer d'autres informations comme les reviews, commentaires ou historique de navigation. De plus, pour améliorer les recommandations, il sera utile de disposer d'une évolution temps-réel

des tags et ressources partagés par les utilisateurs, pour cibler encore mieux leurs besoins.

Remerciements. Ce travail est partiellement financé par le projet franco-tunisien PHC Utique 11G141. Nous remercions les relecteurs anonymes pour leurs remarques constructives.

Références

- BAEZA-YATES R. & RIBEIRO-NETO B. (1999). *Modern Information Retrieval*. Boston, MA, USA : Addison-Wesley Longman Publishing Co., Inc.
- BASILE P., GENDARMI D., LANUBILE F. & SEMERARO G. (2007). Recommending smart tags in a social bookmarking system. In *Bridging the Gap between Semantic Web and Web 2.0*, p. 22–29.
- BELLOGÍN A., CANTADOR I. & CASTELLS P. (2013). A comparative study of heterogeneous item recommendations in social systems. *Inf. Sci.*, **221**, 142–169.
- CERF L., BESSON J., ROBARDET C. & BOULICAUT J.-F. (2009). Closed patterns meet n-ary relations. *ACM TKDD*, **3**, 3 :1–3 :36.
- DAS M., THIRUMURUGANATHAN S., AMER-YAHIA S., DAS G. & YU C. (2012). Who tags what ? an analysis framework. In *Proceedings of PVLDB*, **5**(11), 1567–1578.
- DE MEO P., QUATTRONE G. & URSINO D. (2010). A query expansion and user profile enrichment approach to improve the performance of recommender systems operating on a folksonomy. *User Modeling and User-Adapted Interaction*, **20**(1), 41–86.
- DIEDERICH J. & IOFCIU T. (2006). Finding communities of practice from user profiles based on folksonomies. In *Proceedings of the 1st International Workshop on TEL-CoPs, Crete, Greece*, p. 288–297.
- HERLOCKER J. L., KONSTAN J. A., TERVEEN L. G. & RIEDL J. T. (2004). Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst.*, p. 5–53.
- HU J., WANG B. & TAO Z. (2011). Personalized tag recommendation using social contacts. In *Proc. of Workshop SRS'11, in conjunction with CSCW*, p. 33–40.
- JÄSCHKE R., HOTH A., SCHMITZ C., GANTER B. & STUMME G. (2008). Discovering shared conceptualizations in folksonomies. *Web Semantics.*, **6**, 38–53.
- JÄSCHKE R., MARINHO L., HOTH A., LARS S.-T. & STUM G. (2007). Tag recommendations in folksonomies. In *Proc. of the 11th ECML PKDD, Warsaw, Poland*, p. 506–514.
- JELASSI M. N., BEN YAHIA S. & MEPHU NGUIFO E. (2013). A personalized recommender system based on users' information in folksonomies. In *Proc. of the 5th Intl. Workshop on Web Intelligence & Communities WI&C at 22nd Intl. WWW conf, Rio de Janeiro, May*, p. 1215–1224.
- JELASSI M. N., BEN YAHIA S. & MEPHU NGUIFO E. (2014). Vers des recommandations plus personnalisées dans les folksonomies. *ArXiv e-prints*.
- KIM H. K., OH H. Y., GU J. C. & KIM J. K. (2011). Commenders : A recommendation procedure for online book communities. *Electron. Commer. Rec. Appl.*, **10**(5), 501–509.
- LAM X. N., VU T., LE T. D. & DUONG A. D. (2008). Addressing cold-start problem in recommendation systems. In *Proc. of the 2Nd ICUIMC*, p. 208–211, New York, NY, USA.
- LANDIA N. & ANAND S. (2009). Personalised tag recommendation. *Recommender Systems & the Social Web, New York, NY, USA*, p. 83–86.
- LIPCZAK M. (2008). Tag recommendation for folksonomies oriented towards individual users. In *Proc. of the ECML/PKDD Discovery Challenge, Antwerp, Belgium*, p. 84–95.
- MIKA P. (2007). Ontologies are us : A unified model of social networks and semantics. *Journal of Web Semantics.*, **5**(1), 5–15.
- QUMSIYEH R. & NG Y.-K. (2012). Predicting the ratings of multimedia items for making personalized recommendations. In *SIGIR'12*, p. 475–484.
- RICCI F., ROKACH L., SHAPIRA B. & KANTOR P. (2011). *Recommender Systems Handbook*. Springer.